Detection of Healthcare-Associated Urinary Tract Infection in Swedish Electronic Health Records

Hideyuki TANUSHI¹,², Maria KVIST³,⁴ and Elda SPARRELID⁵,⁶

¹Dept. of Quality and Patient Safety, Karolinska University Hospital, Sweden
²Dept. of Computer and Systems Sciences (DSV), Stockholm University, Sweden
³Dept. of Learning, Informatics, Management and Ethics (LIME), Karolinska Institutet, Sweden
⁴Dept. of Infectious Diseases, Karolinska University Hospital, Sweden

Abstract. The prevalence of healthcare-associated infections (HAI) stresses the need for automatic surveillance in order to follow the effect of preventive measures. A number of detection systems have been set up for several languages, but none is known for Swedish hospitals. We plan a series of infection type specific programs for detection of HAI in electronic health records at a Swedish university hospital. Also, we aim at detecting HAI for patients entering hospital with HAI from previous care, a task that is not often addressed. This first study aims at surveillance of healthcare-associated urinary tract infections. The created rule-based system depends on acquiring the essential clinical information, and a combination of data and text mining is used. The wide range of diverse clinics with different traditions of documentation poses difficulties for detection. Results from evaluation on 1,867 care episodes from Oncology and Surgery show high precision (0.98), specificity (0.99) and negative predictive value (0.99), but an intermediate recall (0.60). An error analysis of the evaluation is presented and discussed.

Keywords. medical records, clinical text mining, patient safety, hospital acquired infections, computer-assisted surveillance

Introduction

Healthcare-Associated Infections (HAIs) result in about 3.2 million patient cases yearly in Europe [1]. In Sweden, about 8–10% of all patients in hospitals have HAI and it contributes to approximately 1,500 deaths per year [2].

It has been estimated that a significant proportion of all HAIs are avoidable with appropriate procedures [3]. Prevention programs are a challenge for all healthcare professionals. Traditional surveillance programs are costly and require extensive manual work, taking time from patient care for clinicians with an already high workload. Computer systems that automatically monitor adverse events, without the need for manual entry, will be important new tools for prevention [4,5] at both the

¹ Corresponding Author: Hideyuki Tanushi, Dept. of Computer and Systems Sciences (DSV), Stockholm University, Forum 100 Kista, Sweden, E-mail: hide-tan@dsv.su.se
organizational level and in direct patient care. A multidisciplinary approach to these problems is necessary.

The aim for this work is to develop better tools for surveillance and prevention of HAI by creating computer systems that automatically examine Swedish electronic health records (EHRs) to detect if an infection is related to the care patients have received. This can be approached by aiming at different specified infections [5] and determining if they are acquired in healthcare, or as a complex approach targeting all HAIs [6]. In this pilot study we focus on detection of healthcare-associated urinary tract infections (HA-UTI) and use data from structured parts of EHR, with addition of some text mining of microbiology reports and the free-text in the records.

1. Background

HAIs are infections that patients acquire during the course of hospital care, or when being examined or receiving treatment in outpatient clinics [2,3]. HA-UTIs are infections of the urinary tract which arise during hospital care and/or as a result of the provision of healthcare or treatment, such as when a patient has a bladder catheter. In the latest multistate US report, HA-UTI was the only HAI that is increasing when other HAIs are responding to preventive measures [3].

1.1. Surveillance of Healthcare-Associated Infections

Hospital-wide surveillance requires control personnel to prospectively and continuously survey all care areas and identify patients with acquired infections. This is, however, very costly. More common are prevalence surveys conducted by counting the number of patients with HAI within a limited time period. In Sweden, point prevalence surveys (PPSs) are performed twice yearly for all hospitalized patients. To reduce burden on personnel, attempts such as selecting patients or areas of the hospital have been made, but this can decrease the reliability of surveillance data. EHR databases enable real-time and prospective computerized HAI surveillance [5].

Computer systems that automatically monitor adverse events such as HAIs have to be adapted to each language and to each EHR system. Most systems use structured data from various hospital databases, but there are systems employing exclusively text mining of the free-text in EHRs, with good results but with a very costly development [6,7]. Freeman et al. [5] present a systematic review of the use of various electronic HAI surveillance systems, most by institutions without employing commercially available software. Most of the studied systems use computer algorithms and data queries for HAI surveillance. Some use Natural Language Processing (NLP) techniques that compare routinely collected data from clinical databases with established HAI surveillance definitions. Other systems create data warehouses for surveillance, in which clinical data are routinely stored to provide surveillance data. In several studies knowledge-based algorithms interpret monitored data by comparing with manual surveillance to classify if the data contains HAI or not. For hospital-wide HAI surveillance, criteria for linked data sources are also applied in several systems. The results vary, but many studies detected HAIs with high recall (mean: 0.90), specificity (mean: 0.86) and negative predictive value (mean: 0.97), but not always so high precision (mean: 0.62).
In California, one third of general acute care hospitals used automated surveillance technology for monitoring HAI in 2008, and the use was positively associated with the depth of implementation of evidence-based practices for infection control [8]. Variations of commercial and custom systems developed at the hospitals were used. MONI-ICU is a surveillance system employing commercial products, currently in use at the Vienna General Hospital [9]. It applies infection criteria including fuzzy logic rules to data from patient records, laboratory and administration. The system gave high sensitivity (0.90) and specificity (1.00) for intensive care unit (ICU) patients.

Machine learning techniques have been used for HAI detection for different languages with good results (e.g. [10–13]). Ehrentraut et al. [12] applied machine learning techniques to detect HAI in Swedish patient records. Based on two classifying algorithms, linguistic pre-processing and feature optimizations for the classifiers were applied with best results of 0.89 recall and 0.76 precision.

### 1.1.1. Automated HA-UTI Surveillance

Several automated HAI surveillance systems are developed for UTI based on definitions of nosocomial infections [5]. However, most of them apply their algorithms only for detection of HAIs that occur after 48 h of hospitalization. Bouam et al. [14] developed an automated HAI surveillance system comprising a series of computer programs that uses test results from microbiology and administrative data from a university-affiliated hospital in France (Table 1). Another French study [15] used four data sources individually or in combination: microbiology, drug prescriptions, medico-administrative and textual discharge summaries, showing importance of microbiology reports. In Taiwan, a HA-UTI surveillance system was created by data mining of six clinical variables [16]. A surveillance tool for catheter associated UTI (CA-UTI) was created utilizing data on fever, urinalysis findings, and urine culture quantitative criteria [17]. At Stanford Hospital, another CA-UTI surveillance tool was found to reduce the need for specialist validation with 97% [18].

<table>
<thead>
<tr>
<th>Group</th>
<th>Infection</th>
<th>Recall</th>
<th>Precision</th>
<th>Specificity</th>
<th>Negative Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bouam et al. (2003) [14]</td>
<td>UTI</td>
<td>0.95</td>
<td>1.00</td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>Bozbid et al. (2011) [15]</td>
<td>UTI</td>
<td>0.98</td>
<td>0.18</td>
<td>0.59</td>
<td>0.998</td>
</tr>
<tr>
<td>Lo et al. (2013) [16]</td>
<td>UTI</td>
<td>1.00</td>
<td>—</td>
<td>0.95</td>
<td>—</td>
</tr>
<tr>
<td>Choudhuri et al. (2011) [17]</td>
<td>CA-UTI</td>
<td>0.86</td>
<td>0.94</td>
<td>0.85</td>
<td>0.94</td>
</tr>
</tbody>
</table>

### 2. Material and Methods

#### 2.1. Data

EHRs² were from three different collections (Table 2) from Karolinska University Hospital, a teaching hospital serving the greater Stockholm area. All records were de-identified. Sections used from the records include the free-text notes by physicians and nurses, lists of diagnoses, drugs and treatment codes, test results from microbiological analyses of blood and urine (in free-text), temperature charts, and administrative records of age, gender, and times of admission and discharge. A care episode was

---

² Ethical permission by the Ethical review board in Stockholm (2012/1838-31/3).
defined as a sequence of days at the hospital, and regarded as a single care episode even if the patients were moved between different departments. Also, if patients returned to hospital within 24 hours of discharge, it was regarded as one care episode. For this study, we included hospitalizations exceeding two days.

<table>
<thead>
<tr>
<th>Table 2. Electronic health records used in the study.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development set 1</td>
</tr>
<tr>
<td>Evaluation set</td>
</tr>
</tbody>
</table>

Development set 1 – Records were collected from patients with HAI as reported by Karolinska University Hospital in the biannual national PPS of April 2012. The PPS of the entire hospital, including the children’s hospital, resulted in a wide variety of patients, disorders and infections in this development set. In total, we received records for 120 patients for a 5 month period describing multiple hospitalizations for some patients, with and without HAI. All care episodes were assessed for occurrence of infections (HAI or non-HAI) by two senior physicians, one a specialist in infectious diseases.

Development set 2 – A number of records from Rheumatology were marked as HAI/non-HAI at the point of patient discharge. Sixty-six adult patients were chosen, half of them marked as having HAI and the other half with non-HAI. As above, records described an extended time period, including care episodes with and without HAI. Of the 134 care episodes, only the 28 episodes with positive microbiology for UTI were assessed by the two physicians.

Evaluation set – A population of 2,000 adult patients treated at surgical or oncologic clinics was chosen for evaluation from the Stockholm EPR Corpus [19,20]. Hospitalizations with duration 2–14 days (1,195 patients), in a 13-month period for which we had all relevant data types, were extracted for evaluation (Table 2).

2.2. Terminologies/classifications Used

Diagnoses were classified using ICD-10 codes (International Classification of Diseases) [21], in a Swedish translation named ICD-10-SE [22]. For catheter use, KVÄ (Klassifikation av Vårdåtgärder, English: classification of procedures of care) codes were used [22]. Antibiotics were identified by ATC-codes for pharmaceutical products [23]. The use of cytostatic drugs was not recorded in the same EHR system, thus unavailable for the study, and instead codes for administering chemotherapy were used.

2.3. HA-UTI Detecting Tool

The HA-UTI detecting tool was developed with Java programming language. As preprocessing, a list of care episodes including admission and discharge time is created according to the study definition. For each care episode, medical data relevant to HA-UTI (i.e. medical records, drugs, microbiology test results, diagnosis and procedure codes, and body temperature) is collected and stored together with its event time. Each medical data for a care episode is checked if it matches or contains the parameters (e.g. bacteria, antibiotics, etc.) relevant to UTI with a combination of data and text mining depending on type of data. For detection of UTI pathogens, text mining of microbiology reports was performed with exact matching to a list of bacteria. For all
care episodes with positive urine culture findings, the tool defines if the care episode includes a HA-UTI according to the criteria (Figure 1 and Table 3) by checking event time of matched medical data.

2.4. Error Analysis

An error analysis was performed after each completed iteration on the development sets, with subsequent program improvement. A final error analysis was performed on the evaluation set. For the first development set, all care episodes were assessed to find reasons for erroneous decisions by the tool. For the second development set and the evaluation set, only care episodes with positive urine cultures were assessed for errors. Assessments were made by two physicians.

3. Results

Several data types were tested for their use in an algorithm for HA-UTI detection, as described below. After choosing data types according to the definition of HA-UTI, and identifying the different fashions these were documented at different wards, the system was run on the development sets in an iterative fashion with adjustments to the program after each iteration. Thus, the system was a result of this process.

Diagnoses are described in the ICD-code system. Unfortunately, there are no codes for HAI. There is a possibility for the healthcare provider to add a code indicating a complication to given care to the code for infection. In our datasets, this combination was very rarely used. Also, the code for UTI was inconsistently used, and sometimes given without proof of bacteria in the urine. For many patients with a long, complicated hospitalization including ICU care, the occurrence of UTI appeared to be of such minor importance that the coding was neglected. We saw a great difference in the efficiency of coding at different departments. Other researchers have also found ICD coding to be a poor tool for infection control surveillance [24]. Therefore, we chose to not use ICD-codes for UTI as a factor in the algorithm.

During the iterative process, a number of changes and additions were made in order to minimize the number of false positives as well as increasing the number of true positives. The number of bacteria and antibiotic variants were restricted to minimize erroneous detection of patients with infections of other types. Antibiotics were detected by ATC-codes while the bacteria type was text mined from the Microbiology reports. To avoid erroneous detection of Asymptomatic Bacteriuria (ABU) as UTI, as it is quite common for some patient groups to carry bacteria without an ongoing infectious process, it was necessary to include data indicating a manifest infection. Symptoms registered in the free-text would require extensive text mining including negation and temporal detection for identification. One symptom, fever, was easily retrieved from the structured data. Iterated development showed that detection of fever (> 38 degrees Celsius) before urinary sampling and/or relevant antibiotics treatment, as a sign of this occurrence being perceived as a clinically manifest disorder, was sufficient to exclude ABUs. A special case of ABU was the finding of Enterococcus in urinary cultures. In order to minimize the false positive UTIs, these bacteria had to be present in both culture of urine and blood to be regarded as a true UTI.
3.1. Definition of HA-UTI for This Study

In this study, for detection in health records, we arrived at a HA-UTI definition as follows: The patient must have a urine culture finding of specific urinary tract pathogens, as well as fever and/or relevant antibiotic treatment (AB) within a limited time related to sampling (Figure 1). The sample should be taken after 48 hours of hospitalization, or, if taken within the first 48 hours of care there must be at least one predisposing healthcare event (Table 3). Predisposing healthcare events chosen were: presence of urinary tract catheter within 14 days, chemotherapy treatment within 30 days, previously transplanted patient or hospitalization within 14 days prior to the current one. The definition of the algorithm for HA-UTI can be expressed as:

\[ \text{[micro} \geq 48\text{h AND (fever OR AB)] OR [micro<48h AND (fever OR AB) AND previous care]} \]

Lists for antibiotics relevant for treatment of UTI, ICD-codes for previous organ or bone marrow transplantation, KVÅ treatment codes for catheter treatment or chemotherapy, words and expressions indicating catheter use sought for in free-text, can be found in the appendix³.

![Figure 1. Illustration of temporal relations of parameters chosen for detection of HA-UTI. Positive urinary culture = detection of bacteria in microbiology reports (free-text). Antibiotics = antibiotic treatment for UTI detected by ATC-codes, given within 72 h of sampling of positive urinary culture. Fever = temperature >38 degrees Celsius, recorded within 24 h preceding sampling of positive urinary culture.](image)

3.2. Identifying HA-UTI During the First 48 h of Hospitalization

For UTI occurring within the first 48 h of hospitalization, the major problem is to distinguish the nosocomial infections from the non-nosocomial infections by selecting and detecting the clinically relevant events indicating previous care related to UTI. In a European PPS, 96% of HA-UTI were associated with urinary catheter use [1]. We chose to identify i) patients with a urinary tract catheter/had recently had a urinary tract catheter, ii) patients that recently received chemotherapy (often given polyclinically), iii) had received a transplant (including immunosuppressive drug therapy) or iv) had recently been hospitalized (Table 3).

Table 3. Temporal relations of additional parameters indicating predisposing healthcare event for detection of HA-UTI for the first 48h of hospitalization. ST-PUC = sampling time of positive urinary culture.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Check start</th>
<th>Check end</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous care</td>
<td>Short interval to previous care episode</td>
<td>14d prior to admission</td>
</tr>
<tr>
<td>Catheter</td>
<td>Urinary tract catheter use, detected in free-text or as treatment code</td>
<td>14d prior to admission</td>
</tr>
<tr>
<td>Risk</td>
<td>ICD-10 codes for previous transplantation</td>
<td>14d prior to admission</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Detected by treatment codes</td>
<td>30d prior to admission</td>
</tr>
</tbody>
</table>

³ [Link to Appendix-UTI-tool-140505.pdf](http://dsv.su.se/polopoly_fs/1.176299.1399376908!/menu/standard/file/Appendix-UTI-tool-140505.pdf)
We identified a few previous events which were possible to detect in the structured data, but catheter use had to be text mined from free-text. In this university hospital, there is presently no structured documentation of the use of urinary catheters. Also, few registered the treatment code for catheter use, but more often used the free-text in the EHR. Therefore, a few words indicating urinary catheter were selected for text mining. For chemotherapy, the module of drug dispensation in the regular EHR-system was not used. As we did not have access to the module for cytostatic drugs, we used the treatment code for chemotherapy. However, also this code was inconsistently used.

3.3. Results for Evaluation Set

Since the records used from PPS and Rheumatology were from patients expected to have a high incidence of infections, the datasets were skewed. The system was therefore tested on a broader population of patients from departments of Surgery and Oncology.

The final program was run on a data set of 1,867 care episodes. Of these, there were positive urine cultures in 89 care episodes and HA-UTI was found by the program in 43 cases (2.3% of all care episodes). This was compared to the assessments made by physicians, finding HA-UTI in 70 care episodes (3.7% of all care episodes). There was only one false positive care episode, resulting in a high precision of 0.98, a specificity of 0.999 and a negative predictive value of 0.99. There were, however, several false negative cases, resulting in a recall (=sensitivity) of 0.60.

Table 4. Results from development and from the final evaluation. F-score is used as a means of averaging recall and precision.

<table>
<thead>
<tr>
<th>Data</th>
<th>Recall</th>
<th>Precision</th>
<th>F-score</th>
<th>Specificity</th>
<th>Negative Predictive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development Set 1</td>
<td>0.80</td>
<td>0.87</td>
<td>0.83</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>Development Set 2</td>
<td>0.57</td>
<td>0.73</td>
<td>0.64</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>Evaluation Set</td>
<td>0.60</td>
<td>0.98</td>
<td>0.74</td>
<td>0.999</td>
<td>0.98</td>
</tr>
</tbody>
</table>

3.3.1. Error Analysis of Evaluation Set

The error analysis for the final evaluation showed that the false negative care episodes were for the most part found among patients entering hospital with an ongoing HA-UTI from previous care, i.e. with a positive urine culture in the first 48 hours of the care episode. Several patients experiencing the onset of infection in their homes were given the advice on telephone counselling to take Paracetamol to reduce fever and this may have affected the algorithms ability to detect HA-UTI.

The errors mainly depended on two problems; finding all instances of urinary tract catheter use and all instances of preceding chemotherapy. Of the false negatives, ten were missed due to lack of certain catheter words, e.g. *Brickerblåsa* (English: Bricker bladder), or due to alternate spellings of listed catheter words. Another eight were deemed negative as preceding chemotherapy was missed. There were also two cases of HA-UTI with positive urine culture without antibiotic treatment as the patients were moved to geriatric hospital before the results of culture was reported, and were given antibiotics there. There were two cases of CA-UTI where the patients died shortly after urinary sampling and no antibiotic treatment had been initialized. Furthermore, in one case a polyclinical biopsy of the prostate preceded a HA-UTI, but the code for that particular operation was lacking from the list of predisposing events.
The list of antibiotics used was focused on compounds directed to urinary tract pathogens in order to minimize detection of patients with other infections. However, sometimes broader antibiotics had been used, resulting in undetected UTI.

4. Discussion

This study shows the feasibility of developing a system for detection of HA-UTI in Swedish EHRs. The study focused on reaching a high specificity, as false alarms could rapidly discourage clinicians from accepting such a surveillance tool. At present, the level of recall is not sufficient for clinical use but the error analysis gave insights to further improvement of the system.

The developed system solved the problem of distinguishing nosocomial from non-nosocomial UTI for the first two days of hospitalization. However, the recall suffered from an incomplete list for catheter words. Solely by adding the lacking words, recall would increase by 14 percentage units (resulting in recall 0.74, F-score 0.84). With larger sets of EHRs, factors such as spelling variations can be accounted for. It has been shown that for text mining the detection of HAI is directly related to the number of iterations performed [7].

For establishing clinical symptoms for UTI, and to distinguish these infections from ABU, further work with text mining of patient reported symptoms in the free-text is needed. One study found limited documentation of subjective symptoms such as urgency, dysuria or pain, compared to the frequent registration of fever. However, for the patients without fever, subjective symptoms were found in a majority of cases [17]. In our study, fever was not a frequent marker for clinically significant UTI and therefore text mining of patient reported symptoms may be of significance. Also studies of CA-UTI has noted the need for text mining to collect relevant information not found in structured parts of ICU [16,18].

The system reaches a performance in line with other studies for specificity, negative predictive value and precision (Table 1). The lower results for recall can in part be attributed to differences in datasets, as previous systems (i.e. [14,15]) were developed for ICU-patients, i.e. patients with a very high frequency of infections. In France, ICU-patients had a 22% rate of HAI [15]. Compared to studies of patients with a known urinary catheter [17,18], our system has a harder task detecting a wider range of HA-UTI not restricted to CA-UTI. Also, we aimed at detection of HAI also for the first 48 h of hospitalization, i.e. patients presenting with HAI acquired in previous care. This can explain lower results than the studies presented in [14–16].

False positive detection could arise when there was a presence of UTI-pathogens in the urine samples without symptoms of an infection, i.e. not UTI but ABU, and a coincident infection of different kind giving rise to fever or treatment with antibiotics. In these instances, treatment and symptoms from other infections (HAI or non-HAI) than UTI complicate the picture and interfere with the program. This was previously seen by other researchers [17], and was seen more in our skewed development datasets than in the evaluation set. If the system is extended to take into account other HAIs in an infection specific pipeline of HAI programs, it may possibly give a better picture of HAI even if it sometimes is hard to determine which infection is the most dominant for the clinical state.
4.1. Limitations

Important limitations for the text mining was that i) we did not make any exhaustive study of terms used to indicate catheter use, and ii) negation detection was not used but it was assumed that the mentioning of a urinary catheter indicated presence or removal of such a device. The skewed datasets were useful for development but may have affected the algorithm. There was incomplete information on chemotherapy in the data, as many of these drugs were registered in a separate computer system, not the regular drug module of the health record system.

4.2. Future Research

Many patients have several caretakers for different disorders, and can also be transferred between different hospitals during the same episode of disease. Connecting multiple healthcare providers in a region to the same surveillance system, would allow conclusions of HAI-scenarios at a regional organizational level without breach of confidentiality as by manual handling of patient records.

Several of the patients that had recently received chemotherapy and presented with a HA-UTI at admission to the hospital were neutropenic. We aim for better recall with addition of more parameters such as leukopenia or other risk factors predisposing for infection. For the next step in the development of this tool, we will include more expressions for catheters, more diagnostic codes for risks, and add ATC-codes for some cytostatic drugs that predispose for HAI. It would also be favourable to gain direct access to databases some of which now do not connect to the main electronic healthcare system, e.g. chemotherapy drug modules.

5. Conclusion

It is feasible to detect HA-UTI in a Swedish EHR system with a combination of data and text mining, also for the first 48 h of hospitalization. Detection of the temporality of events is of essence. The program can separate HA-UTI from non-nosocomial UTI, but not HA-UTI from ABU if the patient has a concomitant antibiotic treatment for a simultaneous infection of another type. The process of completing this program illustrates the need of attention to detail. The results will not be satisfactory for clinical use if the algorithm focus on just a few coarse factors. Surveillance systems must be developed in close cooperation with the clinic and take into consideration the many different traditions of documentation at different workplaces. The study illustrates the need for structured registration in EHR. Structured documentation of catheter use would improve detection, as would also the possibility to follow the patient at several healthcare providers in the county.

Acknowledgements

The research was supported by grants from the Vårdal foundation and from Vinnova (Swedish Governmental Agency for Innovation Systems, grant 2012-01252).
References